Deep Rule-Based Classifier for Action Recognition using Temporal and Spatial Information with 3D Skeleton Joints

THE DEPARTMENT OF
COMPUTER SCIENCE
& ENGINEERING
計算機科學及工程學系

Wong Yu Ning
Advised by Professor Yeung Sai Kit, Supported by ATTA Technologies

INTRODUCTION

One major component of sports game is the accurate recognition of human actions to assess whether players are executing the correct motions. Using traditional rule-based action recognition, however, is too time-consuming and cannot handle subtle variation of human action.

We proposed a video-based multi-action classifier that may assist sports game implemetation. To evaluate the performance of the classifier, we train the classifier to classify 6 actions, namely high knee, jumping jack, lunge, plank, punching and squat.
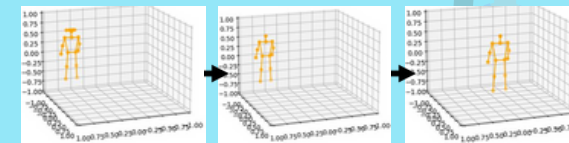
METHODOLOGY

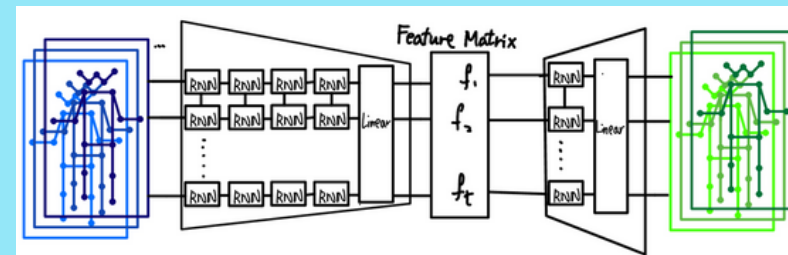1. Use ZED camera to film videos with 3D depth maps

2. Combine YOLO and ZED skeleton estimation module to track subject's actions in 3D space.
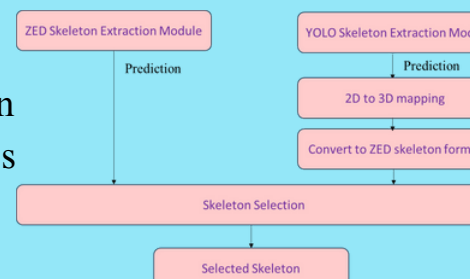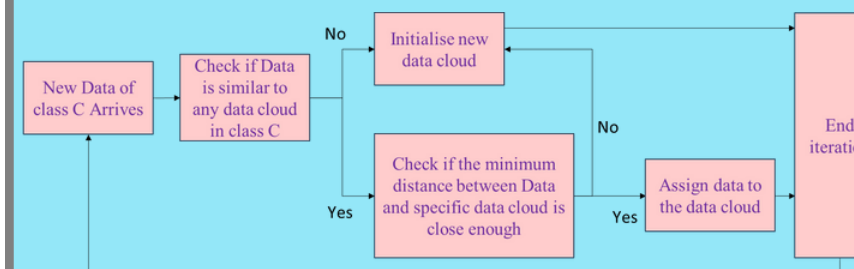

3. Normalise skeleton and remove eyes and ears joints from skeleton sequences.

4. Train a 4-layer autoencoder to extract features from preprocessed skeleton sequences.

5. Train a fuzzy rule-based classifier.

6. Classification:
When test data arrives, the confidence for each class is computed by evaluating the similarity between the data clouds and the test data.

EVALUATION

Accuracy Table of classifier using different percentage of data as training data

Preprocess setup	Optimal Encoded features	Percentage of Data used				
		50%	40%	30%	20%	10%
Original Skeleton	25	0.9986	0.9977	0.9990	0.9983	0.9915
Normalised Skeleton	50	1.0	0.9989	0.9990	0.9983	0.9969
Selected Skeleton	25	1.0	0.9977	0.9960	0.9983	0.9916
Normalised + Selected Skeleton	40	1.0	0.9989	0.9990	0.9990	0.9969
Joint Angles	15	0.9834	0.9839	0.9862	0.9810	0.9531
Normalised + Angles	15	0.9834	0.9805	0.9852	0.9758	0.9470
Selected + Angles	10	0.9793	0.9747	0.9743	0.9706	0.9255
Norm+ Selected+ Angles	10	0.9827	0.9816	0.9822	0.9706	0.9324

Best Classifier

ACTION CLASSIFICATION